Stochastic Processes and their Applications 47 (1993) 113–117
North-Holland

On the Central Limit Theorem for an ergodic Markov chain

K.S. Chan
Department of Statistics and Actuarial Science, The University of Iowa, Iowa City, USA

Received 17 February 1992
Revised 15 September 1992

A simple sufficient condition for the Central Limit Theorem for functionals of Harris ergodic Markov chains is derived. The result is illustrated with an example taken from non-linear time series analysis.

central limit theorem * drift criterion * ergodic Markov chain * non-linear time series model

1. Introduction

Many time series models used in statistical analysis have Markovian representations so that the observations are time-invariant functionals of some possibly unobservable Markov chains. When a time series model has a Markovian representation, its probabilistic properties can be analyzed via the underlying Markov chain. Recently, the theory of Harris ergodic chains has arisen as a new tool to analyze the stationarity of certain time series models. See Tong (1990). Here, we focus on the problem of the CLT for these Harris ergodic models.

In Theorem 7.6 of Nummelin (1984), a sufficient condition for the CLT for a Harris ergodic Markov chain is stated in terms of so-called f-regularity. However, this condition is difficult to verify in practice. Our purpose is to present a criterion in terms of a drift condition (Lyapunov–Foster–Tweedie condition) for the f-regularity and to state a version of the CLT using this criterion. Section 2 contains the main result. In Section 3, we illustrate the criterion via a class of non-linear time series models, the SETAR(2; 1, 1) models.

2. Main result

We first introduce some notations and definitions. Let (X_n) be a Markov chain taking values on a measurable space $(E, \mathscr{E})$ which is assumed to be countably

Correspondence to: Dr K.S. Chan, Department of Statistics and Actuarial Sciences, The University of Iowa, Iowa City, IA 52242, USA.
Research was partially supported by the NSF grants DMS 90-06464 and DMS 91-18626. I would like to thank Professor Howell Tong and an anonymous referee for much useful advices on the style of writing.

generated. In practice, E is often taken as some Euclidean space and $\mathscr{E}$ consists of all the Borel sets. (X_n) is (Harris) ergodic if there exists a unique invariant probability measure, say π, such that $\|P^n(x,\cdot)-\pi\|\to 0$ as $n\to\infty$ where $P^m(x,\cdot)$ is the mth step transition probability for (X_n) and $\|\cdot\|$ denotes the total variation norm. Henceforth, it is assumed that (X_n) is ergodic and π denotes its invariant probability measure. Let $\mathscr{E}^+=\{A\in\mathscr{E}\colon \pi(A)>0\}$. A function $K: E\times\mathscr{E}\to\bar{\mathbb{R}}^+$ is said to be a kernel if $\forall A\in\mathscr{E}$, $K(\cdot, A)$ is a measurable function, and $\forall x\in E$, $K(x,\cdot)$ is a (non-negative) measure. The transition probability $P(x,A)$ is a kernel. For any $B\in\mathscr{E}$, I_B denotes the kernel such that $I_B(x,A)=1$ iff $x\in A\cap B$. For any two kernels K_1 and K_2, the product K_1K_2 is the kernel defined by

$$K_1K_2(x,A)=\int K_2(y,A)K_1(x,\mathrm{d}y).$$

Define

$$G_Bf(x)=E_x\left(\sum_{i=0}^{T_B} f(X_i)\right) \quad\text{and}\quad U_Bf(x)=E_x\left(\sum_{i=1}^{S_B} f(X_i)\right)$$

where $E_x(\cdot)$ denotes taking expectation with the initial distribution concentrated at x and S_B (T_B) is the first $i>0$ $(i\geqslant 0)$ that X_i enters B. Let f be a non-negative π-integrable function. $A\in\mathscr{E}^+$ is said to be f-regular if $\forall B\in\mathscr{E}^+$, U_Bf and f are bounded over A. A non-negative measure λ on E is f-regular if $\forall B\in\mathscr{E}^+$, $\lambda f\overset{\text{def}}{=}\int f(x)\lambda(\mathrm{d}x)$ and λU_Bf are finite. A set $C\in\mathscr{E}^+$ is said to be small if $\exists m$ a positive integer, β a constant and ν a probability measure such that, $\forall x\in C$, $P^m(x,\cdot)\geqslant\beta\nu(\cdot)$. A set $B\in\mathscr{E}$ is said to be special if E is 1_B-regular where 1_B is the characteristic function of B. It is shown in Nummelin (1984) that every small set is special. Further characterizations of f-regularity are given by Nummelin (1984). We now state the main result.

Theorem 1. *Let (X_t) be a Harris ergodic Markov chain whose invariant probability measure is denoted by π. Let $Y_t=h(X_t)$ and h is a measurable function. It is also assumed that $E(Y_t)=0$. Suppose $\exists\gamma>0$, A a small set, and a π-square-integrable and non-negative function $g(x)$ such that*

$$g(x) \text{ and } E(g(X_{n+1})\,|\,X_n=x) \text{ are bounded over } A, \tag{2.1}$$

$$g(x)\geqslant\gamma|h(x)|,\quad x\in A, \tag{2.2}$$

$$g(x)\geqslant E(g(X_{n+1})|X_n=x)+\gamma|h(x)|,\quad x\notin A. \tag{2.3}$$

Then, under any initial distribution λ for X_0, the normalized partial sum $\sum_{i=1}^n h(X_i)/\sqrt{n}$ is asymptotically $\mathrm{N}(0,\sigma^2)$ with σ^2 being finite. If the initial distribution λ of X_0 is such that $E_\lambda(g^2(X_0))<\infty$, then $\sigma^2=\lim_{n\to\infty}E_\lambda(\sum_{i=1}^n h(X_i))^2/n$ where $E_\lambda(\cdot)$ denotes taking expectation with λ being the distribution of X_0.

Proof. For simplicity, we write f for $|h|$ and, with no loss of generality, we also assume that γ in the above conditions is equal to 1.

To prove the asymptotic normality result, Theorem 7.6 in Nummelin (1984) shows that it suffices to check that the measure $v(\mathrm{d}x)=f(x)\pi(\mathrm{d}x)$ is f-regular. This is done as follows. It follows from conditions (2.2) and (2.3) and the Balayage Theorem (Theorem 3.1 in Nummelin, 1984) that

$$g \geq I_{A^c}Pg+f \Rightarrow g \geq G_A f \Rightarrow PI_{A^c}g \geq PI_{A^c}G_A f = U_A f. \tag{2.4}$$

Let B be any Borel set with positive π-measure. By formula (3.18) in Nummelin (1984),

$$U_B f = U_A f + U_B I_{A\setminus B} U_A f. \tag{2.5}$$

It follows from (2.1) and (2.4) that $I_{A\setminus B}U_A f \leq k_1 1_A$ for some positive constant k_1. Because A is small and hence a special set, $U_B 1_A$ is a bounded function. Hence, there exists a constant k_2 such that

$$\int (U_B f)(x) f(x)\pi(\mathrm{d}x) \leq \int g(x)f(x)\pi(\mathrm{d}x) + k_2 \int f(x)\pi(\mathrm{d}x). \tag{2.6}$$

It follows from (2.2), (2.3) and the π-square-integrability of g that the right-hand side of (2.6) is finite. Hence, $v(\mathrm{d}x)$ is f-regular.

A formula for σ^2 is given in Nummelin (1984, p. 134). We now prove the last part of the theorem on the alternative expression for σ^2. We first assume that (X_n) has an atom, i.e., $\forall x \in \alpha$, $P(X_n \in A \mid X_{n-1}=x)$ is the same and $\pi(\alpha)>0$. Let $T_\alpha(i)$ be the epoch that α is visited by (X_t) for the ith time. Define $l(n)$ to be the last epoch $\leq n$ that X_i visits α. Then,

$$\sum_{i=0}^{n} h(X_i) = \sum_{i=0}^{T_\alpha \wedge n} h(X_i) + \sum_{i=1}^{l(n)-1} \zeta(i) + \sum_{i=T_\alpha(l(n))+1}^{n} h(X_i), \tag{2.7}$$

where $l \wedge s = \min(l, s)$ and $\zeta(i)=\sum_{j=T_\alpha(i)+1}^{T_\alpha(i+1)} h(X_j)$.

It is shown in the proof of Theorem 7.6 in Nummelin (1984) that $\{\zeta(j), j=1,2,3,\ldots\}$ is i.i.d. with zero mean and finite variance σ_ζ^2. Furthermore, σ^2 of the limiting Normal distribution is equal to $\pi(\alpha)^{-1}\sigma_\zeta^2$ and $l(n)/n \to \pi(\alpha)$ a.s. as $n\to\infty$. Hence, to prove the stated form for σ^2, it suffices to show that the variances of the first and the third terms on the right-hand side of (2.7) are o(n). Now,

$$E_\lambda\left(\sum_{i=0}^{T_\alpha \wedge n} h(X_i)\right)^2 \leq E_\lambda(G_\alpha f)^2. \tag{2.8}$$

By (2.1)-(2.5), $G_\alpha f \leq a+bg$ for some constants a and b. Hence, the above expectation is finite. Next, it can be shown that

$$E_\lambda\left(\sum_{i=T_\alpha(l(n))+1}^{n} h(X_i)\right)^2 \Big/ n$$

$$\leq \sum_{m=1}^{n} E_\alpha\left(\left(\sum_{i=1}^{m} f(X_i)\right)^2 I(S_\alpha \geq m)\right)\Big/ n \tag{2.9}$$

$$\leq \sum_{m=1}^{n} E_\alpha((U_\alpha f)^2 I(S_\alpha \geq m))/n. \tag{2.10}$$

Again, by (2.1)-(2.5), $\exists$ constants a, b such that $U_\alpha f \leqslant a + bg$. Since $E_\alpha(g^2)$ is finite, $E_\alpha((U_\alpha f)^2 I(S_\alpha \geqslant m))$ tends to zero as $m \to \infty$. Hence, the right-hand side of (2.10) tends to zero as $n \to \infty$. In the general case, the chain can be 'split' to create an atom and then the result can be proved similarly. This completes the proof of Theorem 1. $\square$

Remarks. (1) (X_t) is ϕ-irreducible if $\forall x$ and B with $\phi(B) > 0$, $\text{Prob}(X_n \in B$ for some $n \geqslant 0 \mid X_0 = x) > 0$. Suppose E is some Euclidean space with $\mathscr{E}$ consisting of the Borel sets. It is known that if (X_t) is ϕ-irreducible and Feller-continuous, i.e., $\forall$ bounded and continuous function $w(x)$, $E(w(X_{n+1}) \mid X_n = x)$ is continuous in x, then any compact set with non-zero ϕ-measure is small. See Feigin and Tweedie (1985).

(2) Suppose (X_t) is only known to be aperiodic and ϕ-irreducible and conditions (2.1)-(2.3) hold for some non-negative measurable $g(x)$ and $h(x) \equiv 1$, then it is known that (X_t) is ergodic. The modified condition is known as the 'drift' criterion. If the 'drift' criterion is strengthened by replacing $h(x)$ in (2.2) by 1 and $h(x)$ in (2.3) by $g(x)$, then (X_t) is geometrically ergodic. See Tweedie (1983a). Under the latter condition, Tweedie (1983b) further showed that $E(g(X_t)) < \infty$.

(3) Let γ_i be the covariance of $h(X_1)$ and $h(X_{1+i})$ under the stationary distribution. Suppose that $\sum_{i=0}^{\infty} |\gamma_i| < \infty$. Then the formula of σ^2 implies that $\sigma^2 = \gamma_0 + 2\sum_{i=1}^{\infty} \gamma_i$.

(4) From the proof of the theorem, it can be seen that conditions (2.1)-(2.3) imply that A is $|h|$-regular. It can be similarly shown that a σ-finite measure, λ, is $|h|$-regular if, in addition to (2.1)-(2.3), $\int g(x)\lambda(\mathrm{d}x) < \infty$. See Chan (1991) for a detailed proof.

3. Example

Consider the following simple SETAR(2; 1, 1) model:

$$X_t = \begin{cases} X_{t-1} + \phi_{10} + a_t, & \text{if } X_{t-1} \leqslant 0, \\ X_{t-1} - \phi_{20} + a_t, & \text{if } X_{t-1} > 0, \end{cases} \tag{3.1}$$

where (a_t) is i.i.d. $N(0, 1)$ and a_t is independent of $X_{t-1}, X_{t-2}, \ldots, X_0$. Here, ϕ_{10} and ϕ_{20} are two positive constants. The above model is a special case of the SETAR model which is an interesting class of non-linear time series models. The SETAR model is capable of exhibiting many non-linear phenomena such as limit cycle, strange attractor, ..., etc. See Tong (1990). It is shown in Chan et al. (1985) that (X_t) satisfying (1.1) is ergodic and compact sets with positive Lebesgue measure are small.

We now show that the CLT holds for (X_t). Here, $h(x) = x - \mu$ where $\mu = E(X_t)$. We try $g(x) = x^2 + 1$ and $K = [-r, r]$ for some $r > 0$ which is to be determined later. Now, for $x \leqslant 0$,

$$E(g(X_{n-1}) \mid X_n = x) = x^2 + 2\phi_{10}x + 2 + \phi_{10}^2. \tag{3.2}$$

Together with a similar result for the case when $x > 0$, conditions (2.1)-(2.3) can be verified to hold for some γ sufficiently small and some r sufficiently large. It remains to show that $g(x)$ is square integrable w.r.t. the invariant distribution. This can be done by invoking point (2) of the Remarks in Section 2 and consider, e.g., x^6+1. We note that the Normality assumption for the noise can be substantially weakened. For example, the CLT would continue to hold if the noise has density positive everywhere and has finite sixth moment. The above example demonstrates that Theorem 1 can be effectively used to establish the validity of the CLT for functionals of an ergodic Markov chain. For further applications to non-linear time series models, see Chan (1991).

References

K.S. Chan, Some recurrence properties of an ergodic Markov chain with applications to CLT for functionals of non-linear time series, Tech. Rep. No. 295, Dept. of Statist., Univ. of Chicago. (Chicago, IL, 1991).

K.S. Chan, J.D. Petruccelli, H. Tong and S.W. Woolford, A multiple-threshold AR(1) model, J. Appl. Probab. 22 (1985) 267-279.

P.D. Feigin and R.L. Tweedie, Random coefficient autoregressive processes — a Markov chain analysis of stationarity and finiteness of moments, J. Time Ser. Anal. 6 (1985) 1-14.

E. Nummelin, General Irreducible Markov Chains and Non-negative Operators (Cambridge Univ. Press, Cambridge, 1984).

R.L. Tweedie, Criteria for rates of convergence of Markov chains, with applications to queueing and storage theory, in: J.F.C. Kingman and G.E.H. Reuter, eds., Probability, Statistics and Analysis (Cambridge Univ. Press, Cambridge, 1983a) pp. 260-276.

R.L. Tweedie, The existence of moments for stationary Markov chains, J. Appl. Probab. 20 (1983b) 191-196.

H. Tong, Non-linear Time Series: A Dynamical System Approach (Clarendon Press, Oxford, 1990).